Copyright Notice Addendum

On 15 OCT 2002 The Hollis Group, Inc. submitted this presentation to the US FDA for inclusion in Public Docket 00D-1539, Electronic Records & Electronic Signatures, Retention of Electronic Records. Hollis grants US FDA unlimited rights to copy, distribute, or display this work or an portion thereof, provided attribution to The Hollis Group, Inc. is retained and displayed in the presentation.

Hollis hereby grants any person the right to physically or electronically (I.e. via downloading) copy this presentation for their personal use, provided attribution to The Hollis Group, Inc. is retained and displayed in the presentation.

Hollis specifically prohibits anyone, other than US FDA, from re-copying, re-distributing, re-displaying, or including in an Internet or Intranet website, this presentation, or any portion thereof. In other words, you are allowed to have and view this presentation. If someone else wants it, send them to The Hollis Group, Inc. or to the FDA Docket 00D-1539 to get their own!

Note that this copyright addendum supercedes the copyright notice included in the following presentation, and grants you additional, limited rights to this work.

Copies may be obtained from The Hollis Group, Inc., Station Square Two, Suite 105, Paoli, PA, 19301, (www.hollisgroup.com) or from the Dockets Management Branch (HFA-305), Food and Drug Administration, 5630 Fishers Lane, room 1061, Rockville, MD 20852. The office is open to the public between 9 a.m. and 4 p.m., Monday through Friday. (Note that Hollis and FDA may charge nominal copying and mailing fees for physical copies of this presentation.)

Hollis will prosecute violations of this copyright to the fullest extent of the law.

The Challenge of Long-term Archiving of Electronic Raw Data and Electronic Clinical Data (Part 2)

"What I really need is a droid that speaks the binary language of moisture 'vaporators."

Owen Skywalker

Copyright Notice

The Hollis Group holds and retains exclusive copyright 2001 to these materials. All copying of these materials is strictly prohibited.

These materials have been prepared for the Hollis / Industry Coalition meeting of 17 SEP 2001, and have been distributed electronically in a password-restricted format to attendees of the that meeting. If you are not a meeting attendee, it is a copyright violation for you to be using this electronic file.

Hollis grants permission for each attendee to print one copy of these materials for their personal use.

Reviewing the Agenda…

- E-Archiving requirements
  - State of the industry
  - Regulatory requirements
- E-Archiving architectural challenges
  - Data diversity and obsolescence
  - Maintaining a chain of custody
  - All the angles: tech, legal, RA, QA
  - Currently postulated architectures
- Facilitated discussion

Oh, We Forgot…

- **Encouragement from the U.S. Congress**

- **18 USC 1001 - False information**
- **18 USC 1341 - Mail fraud**
- **18 USC 1343 - Wire fraud**
- **18 USC 1905 - Leaking information**
- **21 USC 331 - Prohibited acts (U.S. Food Drug, & Cosmetic Act)**

The Archiving Dilemma

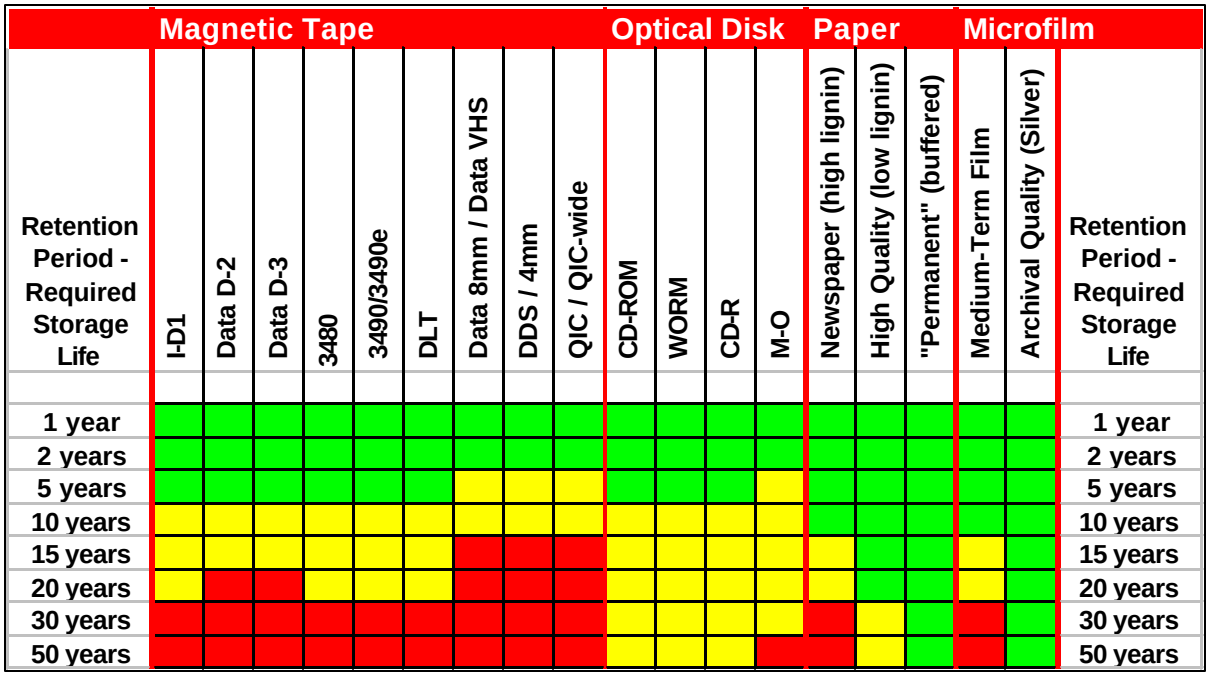
- **We MUST save the records / data for 5 – 30 years, with a chain of custody, and have it available for inspection on demand**
- **Despite decades of research, there are no "magic formats" that are common among broad classes of "live" data / record types**
- **It is not likely that vendors will abandon proprietary formats anytime soon**
- **"Reduced" or static-graphic formats strip the data of key analysis attributes**

A Requirements Example: Clinical Study Data

- We need to make sure that there is no loss or corruption of data during the transfer or migration of data to archives.
- We also need to preserve the attributablity and irrefutability of the data through the transfer and within the receiving system.
- Lastly, we need to provide a mechanism to view the data (and, possibly some of its context) “on demand” for inspections.

The Easy Part: Physical Formats

Media Life Expectancy (LE)
For storage at 20°C (68°F) & 40% RH



Source: National Media Labs, 1994

The Hard Part: Logical Record Formats

- Web pages in various “flavors”
  - Electronic case report forms
  - E-source directly from patients
- Forms in various formats
  - Records from electronic patient diaries
  - Records and reports from adverse event reporting systems
- Protocols, SOP’s, and any other word-processor “documents”
- HTML, XML, .PDF, etc.

“Static” Formats

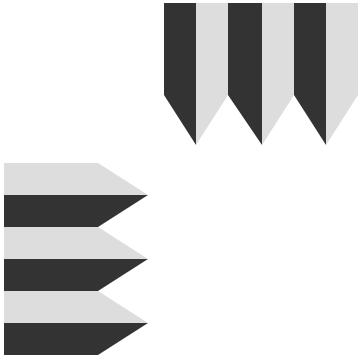
- A number of “standard” record formats exist that are said to be “universal”
  - Many have been around for quite a while
- Reducing a record to such a display format usually strips the metadata and corrupts the chain of evidence.
  - This definitely happens when “database controls” are used for e-signatures
- The longevity of these formats is the subject of considerable debate

The Impossible Part: Operational Arrangements

- Raw data from wearable instruments
- Raw data from clinical instruments
- Data from laboratory instruments
- Data from process control systems
- In other words, e-records that require an “operational arrangement” to be read
- Operational Arrangement – The computer, software, setup parameters, documentation, procedures, and skill needed to run it all

The Breathtaking Pace of Technological Change

- Application software – 1 yr.
- Database software – 2 yr.
- Operating system – 3 yr.
- Compatible hardware – 5 yr.
- Networking standards – 7 yr.
- Computer architectures – 20 yr.

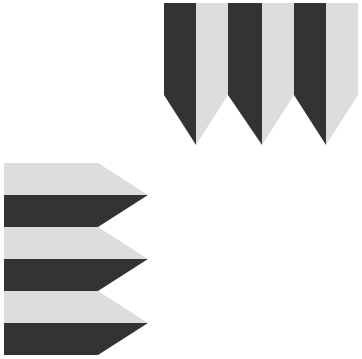


The Inevitability of Data Obsolescence

- Benevolent incompetence is common but has a low impact per incident
- Malice is quite unusual but typically results in serious damage
- Disasters are very rare but they are literally devastating
- Data obsolescence has a probability approaching 100% and a scope of 100%

The Archiving Dilemma

- **We MUST save the records / data for 5 – 30 years, with a chain of custody, and have it available for inspection on demand**
- **“Reduced” or static-graphic formats strip the data of key analysis attributes**
- **Despite decades of research, there are no “magic formats” that are common among broad classes of “live” data / record types**
- **It is not likely that vendors will abandon proprietary formats anytime soon**



The Archiving Solutions

Five basic types of solution:

1. **Convert original formats as required**
2. **Reduce original formats to static formats**
3. **Keep the original operational arrangement**
4. **Build an emulator for the original operational arrangement**
5. **Build a viewer for the original format**

1. Convert Original Formats as Required

- The “constant conversion” solution is prohibitively expensive.
  - But it’s what almost everyone is doing today.
- Conversions typically cost 15% - 20% of the original data acquisition costs.
  - This assumes that the validation goes smoothly.
- The conversion itself is a serious threat to confidentiality, integrity, and availability.
  - Migrating the chain of custody is no picnic, either.

2. Reduce Original Formats to Static Formats

- The “least common data denominator” includes a pretty large fallacy.
- The false promise here is that the “universal” format will be around for a long enough time.
  - SGML, PostScript I, LU 6.2, TTY, anyone?
- In addition, these conversions have all the same conversion challenges as solution 1.

3. Keep the Original Operational Arrangement

- The “Boneyard” concept really appeals to the instructor’s computer packrat side.
  - His wife, on the other hand…
- This solution has the lowest up-front cost.
- Storage costs for the equipment aren’t that bad, either.
- The tough part is hanging on to enough of us old fogey engineers to keep yer geer runnin’!
  - “Missy, I remember a file we put up back in ’02…”

4. Build an Emulator for the Original Operational Arrangement

- An “emulation station” is a very advanced software development project.
  - The validation is even more challenging.
- Since the environment runs original software, there are complicated licensing issues.
  - Translation: Costly
- Under the new modifications to the UCC, it may be possible for software vendors to legally block this solution.

5. Build a Viewer for the Original Format

- **“Read me, Seymore!” requires writing separate viewers for each data format.**
- **It is achievable, but by no means easy.**
- **The industry will need somewhere on the order of 1100 – 1300 viewers.**
- **Someone will have to take the responsibility to write them AND maintain them.**

Let's Run All Five By an Example: - Study Data Transfer

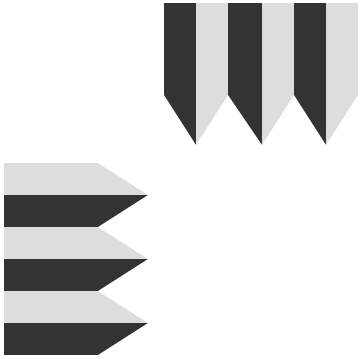
- Transferring study data presents two challenges:
  - Preserving data fidelity and integrity
  - Maintaining a chain of custody
- We need to make sure that there is no loss or corruption of data during the transfer.
- We also need to preserve the attributability and irrefutability of the data through the transfer and within the receiving system.
- Simultaneously achieving these is not easy.

The Legal Angle:

- **Signatures**
  - **Electronic and digital**
- **Certified copies**
  - **How to, who can, what's saved**
- **Notaries**
  - **Trusted third parties since the Pharaohs**
- **Evidence**
  - **Finding and fixing the weakest link**
- **Liability**
  - **Yours, mine, and ours**

Signatures, Electronic Signatures, and Digital Signatures

- Any computer-assisted signature must first "legally" qualify as a signature.
- Remember, the signature is the act, not the ink on the paper or the bits on the disk.
- The act must:
  - Be a firm, knowing declaration / attestation
  - Be performed by the signer or attorney
  - Not be coerced (including implied threats)
  - Include informed consent (of all parties)
  - Produce a durable, understandable residue



Legal and Liability

- **INFOSEC best practice requires the timely investigation of incidents**
  - **The use of audit trails and logs**
  - **Seizing evidence and preserving its integrity**
  - **Logical and physical surveillance**
- **INFOSEC best practice also requires the pursuit of incident resolution**
  - **Incident reporting and follow-up**
  - **Involvement of regulatory or constabulary agencies**

Liability (Cont.)

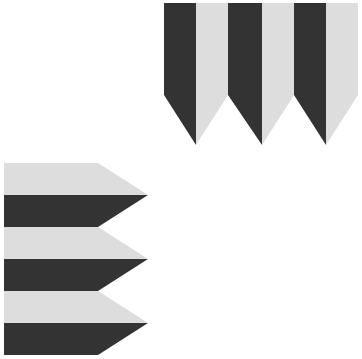
- The investigations procedure will include lost badges, password negligence, attacks, stolen equipment, etc., and incident resolution.
- Since “erroneous” signatures are almost infinitely improbable, unmatched signatures indicate criminal fraud.
  - Investigations will almost always need to involve the legal department.
- We will have to train our people on appropriate use, information ethics, and the personal legal ramifications associated with electronic records and signatures.

Certified Copies and Notaries

- In the United States, the basic mechanism for creating certified copies is the notary public.
- The process must start with a signed or certified document as the source.
- The Notary identifies the person requesting the copy, makes the copy, and affixes a seal to the copy.
- The Notary also records the identity of the requestor, date and time, and a description of the document copied.

Challenges Facing Us With Certified Copies

- The is no national (in the US) or international equivalent of a Notary Public.
- States in the US do recognize notarial seals uniformly for “paper seals.”
- Only 2-3 states recognize electronic notary seals, and the others can’t / don’t recognize out-of-state seals.
- Given this situation, there is much doubt and debate about the legal status of “certified copies” of electronic records required by agencies of the US federal government.



Chain of Custody / Chain of Evidence

- Absolutely trusted people
- Meticulous paperwork
- Counts and reconciliation
- Data and metadata

“Wave” Devices

- Secure delivery of applications and data
- Devices embedded in systems that secure the delivery of programs, data, and authentications
- The “wave” device is a commercial product that is becoming a generic term.
- These devices move the “threshold” of data integrity further into the client environment.

Maintaining a Chain of Custody with "Database Controls"

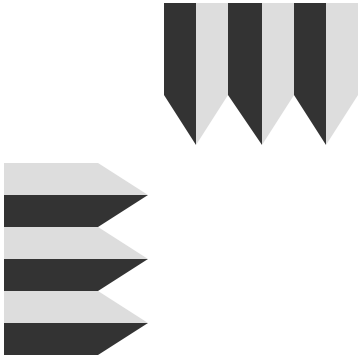
- FDA requires that records be irrefutably associated with their authentic authors.
- In practical terms, this means we must maintain a chain of evidence for e-records.
- 21 CFR 11 relies heavily upon "secure computer-generated, time-stamped audit trails" for this chain of evidence.
- 21 CFR 11 permits "database controls" in lieu of digital signatures and many companies have opted to use "database controls."

The Challenge of Archiving "Database Controls"

- "Database controls" usually consist of transaction journals or application logs of transactions to and from the database.
- If the database engine is not resident in memory, these logs really don't mean much.
- For example, an archived copy of a journal file is just as easy to hack as the original database file.
- So, archiving the journal doesn't increase the security, except to make hacking somewhat more tedious.
- Therefore, "database controls" security measures usually don't "survive" the archiving process.

Physical Security

- We need to plan for the physical security of the facilities, equipment, records, and people.
- The physical security should include at least three “rings” of defenses.
  - Ring 1 - fences, cameras, lighting, guards, etc., to protect the property and personnel
  - Ring 2 - perimeter (building) security of locked doors, alarms, and employee badges
  - Ring 3 - authorization lists, card-key doors, and logs, for critical areas such as data centers
- There should be a formal HVAC, power, and fire protection plan in place.



Access Controls

- Collection of mechanisms
- Directing or restraining influence
- System behavior, content, or use
- Support the system's security goals
  - Confidentiality
  - Integrity
  - Availability
- Include logical and physical access controls

Access Control (Cont.)

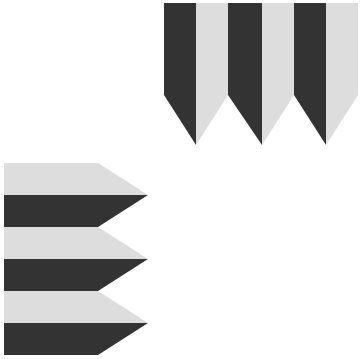
- Physical access controls
  - Didn't we do this in physical security?
  - Yes, but now we focus on controlling "insiders."
  - Here we mean data center access, media handling, equipment configuration management, etc.
- Logical security controls
  - Password / logon system
  - Identification / authentication system
  - Data classification by access type
  - Authorization engine to enforce classifications

Access Control Methods

Access Control Type vs. Arrangement	Physical	Logical
Perimeter	Fences, gates, locks, badges	Identification, login, encryption
Operational	Surveillance, media handling, logbooks	System audits, network monitors, firewalls

Access Control (Cont.)

- Identity verification procedures
- Password and token management
- Logs, journals, alarms, signals, and reports
- The “two checks” rule (i.e., two systems)
- Good access control practices:
  - Rule of least privilege, privilege matrices
    - Job-specific or need-to-know access
  - Separation of duties, segregation of data
  - Ownership, accountability, and reconciliation



Logical Access Controls: Data Classification

- Basically, this is aggregating data by use.
- Classified data is labeled with security (access privilege) attributes.
- The authorization subsystem manages access, using the labeling information.
- Commercial and military schemes exist.
- The basic tasks include labeling, marking, storage, logging, and copying.
- The classification scheme addresses data age, sensitivity, copy limits, useful life, etc.

Access Control Matrix (By Record Type)

System / UserID	Calibration SOP's	Master Records	BOM's	Laboratory Procedures	Raw Material Specifications	Drawings
John Doe	R-W-E D-M-A	R-W-E D-M-A	R-W-E D-M-A	R-W-E D-M-A	R-W-E D-M-A	R-W-E D-M-A
Jane Doe	R-W-E D-M-A	R-W-E D-M-A	R-W-E D-M-A	R-W-E D-M-A	R-W-E D-M-A	R-W-E D-M-A
Joe Smith	R-W-E D-M-A	R-W-E D-M-A	R-W-E D-M-A	R-W-E D-M-A	R-W-E D-M-A	R-W-E D-M-A
Mary Smith	R-W-E D-M-A	R-W-E D-M-A	R-W-E D-M-A	R-W-E D-M-A	R-W-E D-M-A	R-W-E D-M-A

Read
Write
Edit
Delete
Move
Archive

Long-term Retention Formats

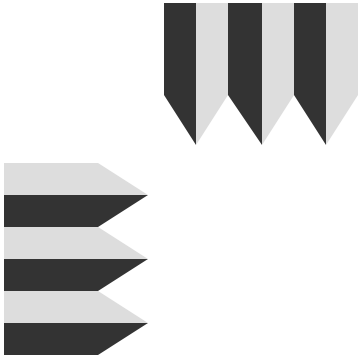
- Physical considerations
  - CD's, tapes, disks, cards (and more exotic ones)
  - information density and footprint cost
  - native degradation (including obsolescence)
  - degradation vs. storage environment
- Logical – static or live
  - Static – images
  - Live – raw data
- And the hardware to read the media

Media Life Expectancy (LE)

For storage at 20°C (68°F) & 40% RH

	Magnetic Tape									Optical Disk				Paper			Microfilm		
Retention Period - Required Storage Life	I-D1	Data D-2	Data D-3	3480	3490/3490e	DLT	Data 8mm / Data VHS	DDS / 4mm	QIC / QIC-wide	CD-ROM	WORM	CD-R	M-O	Newspaper (high lignin)	High Quality (low lignin)	"Permanent" (buffered)	Medium-Term Film	Archival Quality (Silver)	Retention Period - Required Storage Life
1 year																			1 year
2 years																			2 years
5 years																			5 years
10 years																			10 years
15 years																			15 years
20 years																			20 years
30 years																			30 years
50 years																			50 years

Source: National Media Labs, 1994



NAS and SAN

- **NAS – Network attached storage**
  - **Dedicated, pre-configured servers with large integral disk capacities and that use a standard operating system for network interface**
- **SANs – Storage array networks**
  - **Dedicated storage arrays that include proprietary hardware control operating systems internally, management software, and a standard network interfcae, such as TCP/IP**
- **The line between these two kinds of devices is quite gray.**

The RA angle: Records Retention Planning

- We must have a records retention plan.
- The plan includes:
  - The list of retained records
  - The retention schedule
  - The logical and physical retention formats
  - The storage locations and retrieval procedures
  - The restoration procedures
  - The required equipment / software list
  - The chain-of-custody model

Records Destruction

- Paper, tapes, CD-ROM’s and hardware
- Again. Trusted people
- Big legal question: when to destroy
- We need to remove “links” to destroyed data.

System Decommissioning

- There must be a formal procedure for removing components from service.
- The procedure must include an analysis of the effect of the removal on system function and data resident on the component.
- The removal process must include updating configuration drawings and documents
- Decommissioning must include securely deleting all information from the system.

Records Management Audits

- Our standard audit plan needs to include the electronic records management functions and systems at the archive site.
- Particular areas of interest include:
  - Access controls to operational areas
  - Media labeling, handling, and reconciliation
  - Password challenge testing
  - Badge and token procedures
  - Disaster recovery rehearsals
  - Security awareness training

Archive Availability and Restoration Audits

- We must periodically audit and test our ability to restore records in a usable format.
- We must also audit:
  - Records retention schedule compliance
  - Applicable sections of the disaster recovery plan
  - Records destruction procedure compliance
  - Storage facility operation and management
    - Note that the last two will include significant security audit components, such as personnel background checks, reconciliation, and destruction methods.

Lunch!

Questions? Discussion?

Contact Information:

Thomas Quinn,
President
tquinn@hollisgroup.com

Barbara Meserve,
Vice President, QA
bmeserve@hollisgroup.com

The Hollis Group, Inc.
Station Square Two, Suite 109
Paoli, PA 19301

v - 610-889-7350
f - 610-296-2339